DAO Office Note 96-18

Office Note Series on Global Modeling and Data Assimilation

Richard B. Rood, Head
Data Assimilation Office
Goddard Space Flight Center
Greenbelt, Maryland

Data assimilation in the presence of forecast bias

Dick Dee*
Arlindo da Silva†

Data Assimilation Office, Goddard Laboratory for Atmospheres
* *General Sciences Corporation, Laurel, Maryland*
† *Goddard Space Flight Center, Greenbelt, Maryland*

Goddard Space Flight Center
Greenbelt, Maryland 20771
December 1996

Abstract

Statistical analysis methods are generally derived under the assumption that forecast errors are strictly random and zero in the mean. If the short-term forecast, used as the background field in the statistical analysis equation, is in fact biased, so will the resulting analysis be biased. The only way to properly account for bias in a statistical analysis is to do so explicitly, by estimating the forecast bias and then correcting the forecast prior to analysis.

We present a rigorous method for estimating forecast bias by means of data assimilation, based on an unbiased subset of the observing system. The result is a sequential bias estimation and correction algorithm, whose implementation involves existing components of operational statistical analysis systems. The algorithm is designed to perform on-line, in the context of suboptimal data assimilation methods which are based on approximate information about forecast and observation error covariances. The added computational cost of incorporating the algorithm into an operational system roughly amounts to one additional solution of the statistical analysis equation, for a limited number of observations. Off-line forecast bias estimates based on previously produced assimilated data sets can be produced as well, using an existing analysis system.

We show that our sequential bias estimation algorithm fits into a broader theoretical framework provided by the *separate-bias estimation* approach of estimation theory. In this framework the bias parameters are defined rather generally and can be used to describe systematic model errors and observational bias as well. We illustrate the application of on-line forecast bias estimation and correction in a simulated data assimilation experiment with a one-dimensional forced-dissipative shallow-water model. A climate error is introduced into the forecast model via topographic forcing, while random errors are generated by stochastic forcing. In this simple experiment our algorithm is well able to estimate and correct the forecast bias caused by this systematic error, and the climate error in the assimilated data set is virtually eliminated as a result.

Contents

List of Figures

1 Introduction

Atmospheric data assimilation systems combine observational data with a background field, usually consisting of a short-term model forecast, in order to produce accurate and comprehensive analyses of atmospheric fields and parameters. Optimal analysis accuracy, in a proper statistical sense, is obtained when the ensemble means and ensemble covariances of the error fields associated with the model forecasts and with the observations are known and accurately specified. Since these statistics are not generally available, actual implementations of statistical data assimilation algorithms are always suboptimal.

A large portion of the research pertaining to the specification of error statistics in data assimilation systems has concerned *covariance modeling*, which is the development of methods for representing and estimating forecast and observation error covariances. Error statistics required for *optimal interpolation* (OI) are usually estimated from time series of observed-minus-forecast residuals (Rutherford 1972; Hollingsworth and Lönnberg 1986; Lönnberg and Hollingsworth 1986; Daley 1991; Bartello and Mitchell 1992). Advanced statistical data assimilation techniques aim to improve the accuracy of forecast error statistics by taking into account the effect of model dynamics on the evolution of forecast errors (Ghil *et al.* 1981; Dee 1991; Cohn and Todling 1996).

The point of departure in covariance modeling is complete knowledge of the means. Most often it is simply assumed that the forecast model as well as the observing instruments are unbiased; that is, the mean errors are zero or they have been removed. Identification and correction of observational bias is an important component of operational data assimilation systems. Examples include radiation correction procedures for radiosonde observations (Julian 1991), and bias removal schemes for cloud-cleared radiance (Eyre 1992). Some numerical weather prediction centers use 6-hour model forecasts to provide a reference for removing bias from the observations (Baker 1991), at the risk of perpetuating any existing biases in the forecast itself.

The term *forecast bias* is synonymous with *non-zero mean forecast error;* if present, the forecast model is a biased estimator of the actual atmosphere. Forecast bias is due to the presence of systematic errors in the forecast model, such as are caused by incorrect physical parameterizations, numerical dispersion, or faulty boundary conditions. Often the effects of such errors persist for a certain amount of time, and are detected when specific aspects of the model climatology differ from the actual atmospheric climatology as derived from observations. Although it is well known that systematic errors contribute significantly to forecast errors (see, for example, Reynolds *et al.* 1996), the problem of estimating and properly accounting for forecast bias in data assimilation systems has received little attention so far.

Saha (1992) has estimated forecast bias in the U.S. National Centers for Environmental Predictions (NCEP)[1] model by averaging one month of differences between one-day forecasts and the verifying operational analyses. It is not uncommon to evaluate systematic errors in a forecast model by using analyses as a reference (e.g., Takacs and Suarez 1996). The success of this approach obviously depends on the validity of the underlying assumption that the analyses themselves are unbiased. Tenenbaum (1996) has shown by using independent (i.e., not assimilated) aircraft data that analyzed jet stream winds obtained from various operational centers are significantly biased. The likely explanation for this is that the analyses are produced from biased

[1]Formerly the National Meteorological Center (NMC).

forecasts; sparse observations of jet stream winds will, at best, only partially remove this bias. Thus, if forecast bias is a problem, then it is not safe to assume that analyses are unbiased.

The purpose of this article is to present a rigorous, yet practical, method for estimating forecast bias in an atmospheric data assimilation system. The method is fully consistent with the state-space approach of estimation theory, originally presented in the context of atmospheric data assimilation by Ghil *et al.* (1981). This theory requires explicit assumptions on statistics of observation errors and on forecast errors, possibly including unknown systematic (i.e., non-zero mean) components. From these assumptions it is then possible to derive a consistent set of algorithms for estimating forecast bias and for producing unbiased analyses.

The basic assumption we adopt here is that there exists a subset of the observing system for which bias is negligible compared to the forecast bias. In addition, we explicitly define forecast bias as the time-mean (climatological) error in the short-term forecast, and this is the quantity we set out to estimate. We are then able to derive a rigorous sequential forecast bias estimation algorithm, whose implementation involves existing components of statistical data assimilation systems. Consequently one can incorporate forecast bias estimation in an operational system with only minor modifications. The algorithm is designed to perform in the context of suboptimal data assimilation methods in which error covariance information is only approximate. The added computational cost of on-line forecast bias estimation is roughly one additional solution of the statistical analysis equation. Off-line forecast bias estimates can be produced as well, using an existing data assimilation system and stored output from a previous data assimilation run.

To provide our bias estimation algorithm with a firm theoretical footing, we briefly review the so-called *separate-bias estimation* approach of estimation theory. Friedland (1969) formulated the bias estimation problem for a class of linear stochastic-dynamic systems with constant bias parameters, and showed that estimates of these parameters can be obtained separately from the estimates of the dynamic state variables. Other authors subsequently clarified and extended Friedland's formulation (e.g., Tacker and Lee 1972; Mendel 1976; Friedland 1978; Ignagni 1981; Ignagni 1990; Zhou *et al.* 1993). Separate-bias estimation algorithms can be applied more generally to estimate model error parameters and observational bias as well.

The organization of this paper is as follows. In section 2 we discuss forecast errors and their statistics, and show that the usual statistical analysis equation produces biased analyses in the presence of forecast bias. We show in section 3 how forecast bias can be estimated sequentially in a data assimilation system, provided unbiased (or bias-corrected) observations are available. Section 4 contains a concise review of the bias estimation theory originally developed by Friedland (1969), and there we reconcile our approach to forecast bias estimation with this theory. In section 5 we discuss certain practical aspects of forecast bias estimation, for off-line as well as on-line implementations. Here we also describe a simple numerical experiment based on a linear, one-dimensional shallow water model with topographic and stochastic forcing. The climate of the forecast model in this experiment differs from the simulated 'true' climate, and we show that our algorithm successfully corrects this systematic error. We briefly conclude in section 6.

2 Bias and the analysis equation

Here we show that a biased forecast invariably leads to a biased analysis, independently of the weights used in the analysis update. Bias can be reduced by assigning more weight to the observations, but the result will be an increasingly noisy analysis. We also briefly discuss the distinction between ensemble means and time averages. We first define forecast and observation errors and their first- and second-order statistics.

2.1 Forecast and observation errors.

Suppose that the n-vector $\mathbf{w}_k^f$ is a model forecast valid for time t_k, and $\mathbf{w}_k^t$ is the unknown true state of the atmosphere at that time. It is convenient to define both quantities in terms of the same state representation: $\mathbf{w}_k^t$ is an n-vector as well, containing, for example, the true grid-point values or spectral coefficients. The *forecast error* is then simply

$$\boldsymbol{\varepsilon}_k^f \equiv \mathbf{w}_k^f - \mathbf{w}_k^t. \tag{1}$$

For a p_k-vector $\mathbf{w}_k^o$ of measurements generated by a particular instrument at time t_k, the *observation error* is defined by

$$\boldsymbol{\varepsilon}_k^o \equiv \mathbf{w}_k^o - \mathbf{h}_k(\mathbf{w}_k^t). \tag{2}$$

The nonlinear p_k-vector function h_k is the *discrete forward observation operator* (e.g., Cohn 1996), mapping model variables to the data type associated with the instrument.

We introduce the following notation for the forecast error mean and covariance

$$\mathbf{b}_k^f \equiv \langle \boldsymbol{\varepsilon}_k^f \rangle, \qquad \mathbf{P}_k^f \equiv \left\langle (\boldsymbol{\varepsilon}_k^f - \mathbf{b}_k^f)(\boldsymbol{\varepsilon}_k^f - \mathbf{b}_k^f)^T \right\rangle, \tag{3}$$

and for the observation error mean and covariance

$$b_k^o \equiv \langle \boldsymbol{\varepsilon}_k^o \rangle, \qquad \mathbf{R}_k \equiv \left\langle (\boldsymbol{\varepsilon}_k^o - \mathbf{b}_k^o)(\boldsymbol{\varepsilon}_k^o - \mathbf{b}_k^o)^T \right\rangle. \tag{4}$$

Here $\langle \cdot \rangle$ denotes the *ensemble average* or *expectation operator*, whose proper definition involves the joint probability distribution of forecast and observation errors.

In order to simplify the presentation we will assume throughout that observation and forecast errors are uncorrelated:

$$\left\langle (\boldsymbol{\varepsilon}_k^o - \mathbf{b}_k^o)(\boldsymbol{\varepsilon}_k^f - \mathbf{b}_k^f)^T \right\rangle = 0. \tag{5}$$

Removal of this assumption does not introduce any significant complications in what follows.

A forecast $\mathbf{w}_k^f$ is said to be *biased* if the mean forecast error $\mathbf{b}_k^f$ is nonzero; $\mathbf{b}_k^f$ is the *forecast bias*. Similarly, the observations $\mathbf{w}_k^o$ are said to be biased if the mean observation error or *observation bias* $\mathbf{b}_k^o$ is nonzero.

2.2 Ensemble means vs. time averages.

We defined forecast and observation error statistics in terms of *ensemble means:* these are averages over all possible realizations of the errors, weighted by their probability of occurrence. This definition is appropriate since the optimality criteria underlying state estimation algorithms are generally formulated in terms of probability distributions of the stochastic-dynamic state variables (Jazwinski 1970; Cohn 1996). For example, the optimal estimate (in a rather broad sense) of the true atmospheric state $\mathbf{w}_k^t$ given any set W of observations is provided by the conditional (ensemble) mean $\langle \mathbf{w}_k^t | W \rangle$. This quantity is defined in terms of the joint probability distributions of $\mathbf{w}_k^t$ and W.

Note, however, that the ensemble of all possible realizations of the *actual* atmospheric state is different from the ensemble of all possible realizations of the *modeled* atmospheric state, both in concept and in substance. *Ensemble forecasting* (Toth and Kalnay 1993; Houtekamer *et al.* 1996) involves different realizations of model forecasts obtained by perturbing initial conditions and/or selected model parameters; the number of such realizations is limited only by the computing resources at hand. The ensemble of actual atmospheric states, on the other hand, is nothing more than a theoretical device. Only a single member of this ensemble exists, and only this single physical realization of the atmospheric state evolution is in fact observable; all general inferences about the ensemble rely on theory. For example, to assert that properties of the ensemble of actual states can be emulated by generating an ensemble of modeled states involves assumptions on the exact relationship between the model and the real atmosphere.

In practice, first- and second-order forecast and observation error statistics are computed by averaging over time, usually over periods on the order of a month or so (Rutherford 1972; Schlatter 1975; Lorenc 1981; Bartello and Mitchell 1992). Substitution of ensemble means by some other kind of average is, of course, a practical necessity. One could attempt to justify this substitution by assuming ergodicity of the stochastic processes involved, although this would seem to be rather farfetched. We will not further address this issue here but simply keep in mind the practical definition of forecast and observation error statistics in terms of time averages as an alternative to the theoretical definition in terms of ensemble means.

Our notion of *forecast bias* in particular is usually associated with errors that persist for a certain amount of time. Such systematic errors are detectable when they cause specific aspects of the model climatology to differ from the actual atmospheric climatology, as derived from observations.

2.3 The analysis equation in the presence of bias.

If the forecast bias were known, one could compute an unbiased forecast

$$\widetilde{\mathbf{w}}_k^f = \mathbf{w}_k^f - \mathbf{b}_k^f. \tag{6}$$

Similarly,

$$\widetilde{\mathbf{w}}_k^o = \mathbf{w}_k^o - \mathbf{b}_k^o \tag{7}$$

would be a set of unbiased observations. Throughout this paper we will use tildes to indicate that a quantity is either unbiased (in case of an estimate) or that its mean is zero (in case of a random vector).

To simplify the presentation we now assume that the observation operator is linear: $\mathbf{h}_k(\cdot) = \mathbf{H}_k\cdot$ in (2), with $\mathbf{H}_k$ a $p_k \times n$ matrix. The statistical analysis equation which properly accounts for bias is then

$$\widetilde{\mathbf{w}}_k^a = \widetilde{\mathbf{w}}_k^f + \mathbf{K}_k \left[\widetilde{\mathbf{w}}_k^o - \mathbf{H}_k \widetilde{\mathbf{w}}_k^f\right], \tag{8}$$

where $\widetilde{\mathbf{w}}_k^a$ is the analysis at time t_k, and $\mathbf{K}_k$ is an $n \times p_k$ *gain matrix* which takes into account the relative accuracies of forecast and observations. Independently of the specification of this gain, the analysis is an unbiased estimate of the true atmospheric state:

$$\mathbf{b}_k^a \equiv \langle \boldsymbol{\varepsilon}_k^a \rangle = 0, \qquad \boldsymbol{\varepsilon}_k^a \equiv \widetilde{\mathbf{w}}_k^a - \mathbf{w}_k^t. \tag{9}$$

If, in particular,

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k\right]^{-1}, \tag{10}$$

then (8) provides the linear minimum variance estimate of the true atmospheric state given all observations up to and including time t_k (Anderson & Moore 1979, section 5.2).

In operational data assimilation systems the bias terms $\mathbf{b}_k^o, \mathbf{b}_k^f$ are usually unknown and hence neglected. Using $\mathbf{w}_k^o, \mathbf{w}_k^f$ in place of $\tilde{\mathbf{w}}_k^o, \tilde{\mathbf{w}}_k^f$ the analysis equation is

$$\mathbf{w}_k^a = \mathbf{w}_k^f + \mathbf{K}_k \left[\mathbf{w}_k^o - \mathbf{H}_k \mathbf{w}_k^f\right]. \tag{11}$$

Taking the ensemble average and using (6) and (7) implies

$$\mathbf{b}_k^a = \mathbf{b}_k^f + \mathbf{K}_k \left[\mathbf{b}_k^o - \mathbf{H}_k \mathbf{b}_k^f\right], \tag{12}$$

which shows that, for any gain $\mathbf{K}_k$, the analysis is biased unless the forecast as well as the observations happen to be unbiased. Equation 12 also shows that the mean analysis increment (the second term on the right-hand side) does not provide a good estimate of forecast bias, even when $\mathbf{b}_k^o \equiv 0$, since the gain coefficients are generally less than one.

Given an analysis equation of the form (11) in which bias is not explicitly accounted for, it is nevertheless interesting to consider the particular gain $\overline{\mathbf{K}}_k$ which leads to the smallest total (systematic plus random) root-mean-square (rms) analysis error. This is important from a practical point of view since (11) is precisely the equation being solved in operational sequential data assimilation systems. It is not difficult to show that the rms analysis error due to (11) is minimal for

$$\overline{\mathbf{K}}_k = \overline{\mathbf{P}}_k^f \mathbf{H}_k^T \left[\mathbf{H}_k \overline{\mathbf{P}}_k^f \mathbf{H}_k^T + \overline{\mathbf{R}}_k\right]^{-1}, \tag{13}$$

with

$$\overline{\mathbf{P}}_k^f \equiv \left\langle \boldsymbol{\varepsilon}_k^f (\boldsymbol{\varepsilon}_k^f)^T \right\rangle = \mathbf{P}_k^f + \mathbf{b}_k^f (\mathbf{b}_k^f)^T, \tag{14}$$

$$\overline{\mathbf{R}}_k \equiv \left\langle \boldsymbol{\varepsilon}_k^o (\boldsymbol{\varepsilon}_k^o)^T \right\rangle = \mathbf{R}_k + \mathbf{b}_k^o (\mathbf{b}_k^o)^T. \tag{15}$$

The analysis resulting from (11) with $\mathbf{K}_k = \overline{\mathbf{K}}_k$ is still biased, as is true for any gain $\mathbf{K}_k$. An unbiased analysis can be produced only if explicit estimates of forecast bias and observation bias are available.

2.4 A scalar example.

Suppose that w_k^f and w_k^o are both scalars, with

$$b_k^f = \langle \varepsilon_k^f \rangle = b, \qquad P_k^f = \langle (\varepsilon_k^f - b)^2 \rangle = \sigma^2, \tag{16}$$

$$b_k^o = \langle \varepsilon_k^o \rangle = 0, \qquad R_k = \langle (\varepsilon_k^o)^2 \rangle = \sigma^2. \tag{17}$$

Using (8), the optimal analysis is given by

$$w_k^a = \frac{1}{2}(\tilde{w}_k^f + w_k^o) = \frac{1}{2}(w_k^f - b + w_k^o), \tag{18}$$

for which

$$b_k^a = 0, \qquad \langle (\varepsilon_k^a)^2 \rangle = \frac{1}{2}\sigma^2. \tag{19}$$

Ignoring forecast bias as in (12) would give instead

$$w_k^a = \frac{1}{2}(w_k^f + w_k^o), \tag{20}$$

which is biased:

$$b_k^a = \frac{1}{2}b, \qquad \langle (\varepsilon_k^a)^2 \rangle = \frac{1}{4}b^2 + \frac{1}{2}\sigma^2. \tag{21}$$

The magnitude of the mean analysis increment in this case is $b/2$ and would underestimate the forecast bias by a factor of two.

Note that the analysis reduces the bias but does not remove it. Suppose now that $b = \sigma$, i.e. the typical magnitude of the random component of forecast error is equal to that of the systematic component. Increasing the weight of the observation as in (13) then gives

$$w_k^a = \frac{1}{3}(w_k^f + 2w_k^o), \tag{22}$$

which is still biased but has somewhat less total variance:

$$b_k^a = \frac{1}{3}b, \qquad \langle(\varepsilon_k^a)^2\rangle = \frac{1}{9}b^2 + \frac{5}{9}\sigma^2. \tag{23}$$

Drawing the analysis even closer to the observation would further reduce the bias but increase the total analysis error variance, due to the random error component. Figure 1 summarizes this example; it shows the dependence on the weight K of the analysis bias, the standard deviation of the random component of analysis error, and the total expected analysis error if (12) is used. This example shows clearly that, unless bias is explicitly accounted for, it can be reduced only at the expense of increasing the noisiness of the analysis.

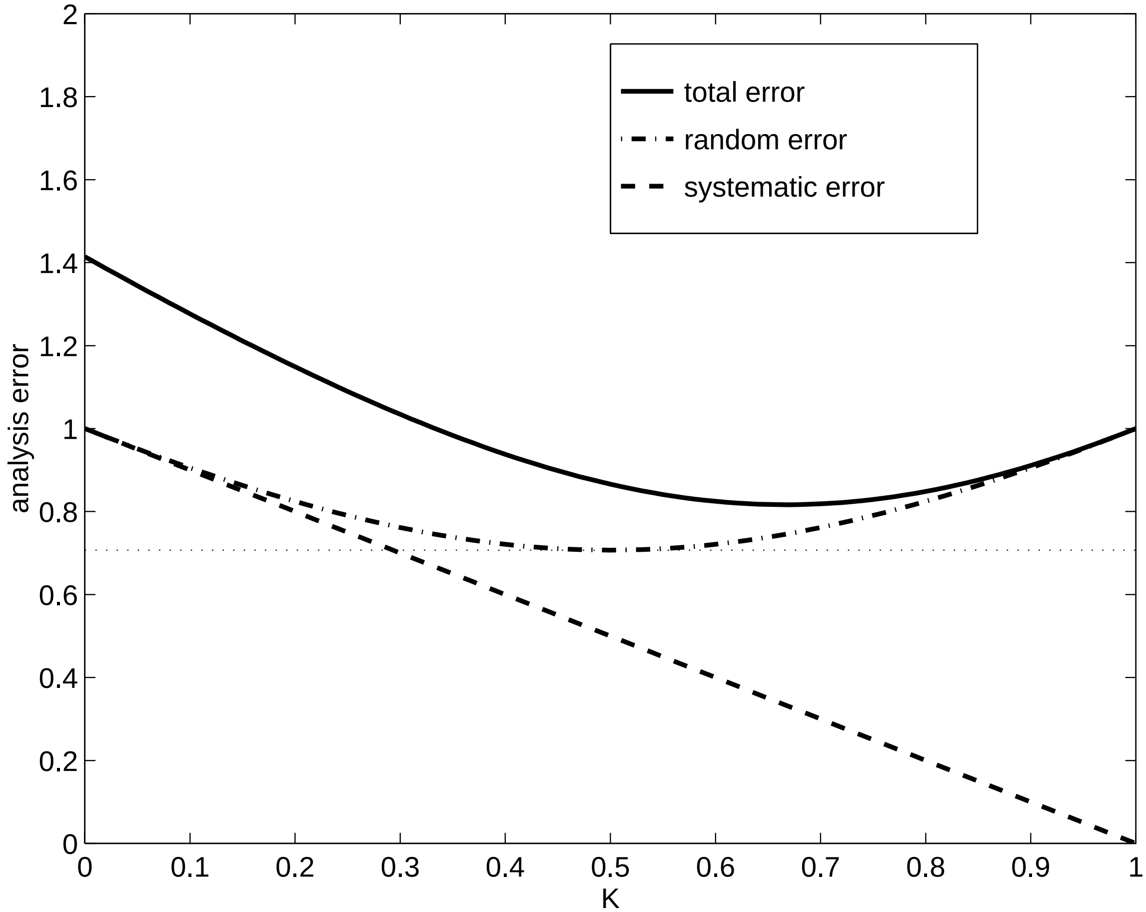

Figure 1: Analysis error as a function of the scalar gain coefficient K, when forecast bias is not explicitly accounted for in the analysis, for the scalar example presented in section 2. The dotted horizontal line indicates the minimum analysis error level obtainable with an unbiased forecast.

3 Sequential bias estimation

Forecast bias can be estimated by comparing forecasts with observations, i.e., from observed-minus-forecast residuals. Without additional information it is not possible to separate the effect of forecast bias on these residuals from that of biased observations. We therefore assume, in this section, that a subset of the observing system is unbiased. This then leads to a sequential estimation algorithm for the time-averaged forecast error. First, we briefly discuss observed-minus-forecast residuals and their first- and second-order statistics.

3.1 Observed-minus-forecast residuals.

The observation operator introduced in (2) is a device for comparing forecasts with observations. The *observed-minus-forecast residuals* defined by

$$\mathbf{v}_k \equiv \mathbf{w}_k^o - \mathbf{h}_k(\mathbf{w}_k^f) \tag{24}$$

are routinely computed in operational data assimilation systems. The residual p_k-vector time series $\{\mathbf{v}_k\}$ is often referred to as the *innovation sequence*, although this terminology is not entirely correct since it presumes optimality of the data assimilation algorithm (Anderson & Moore 1979, section 5.3). In any case, these residuals contain important information about the actual observation and forecast errors, since

$$\mathbf{v}_k \approx \boldsymbol{\varepsilon}_k^o - \mathbf{H}_k \boldsymbol{\varepsilon}_k^f, \tag{25}$$

where the linearized observation operator $\mathbf{H}_k$, a $p_k \times n$-matrix, is defined by

$$\mathbf{H}_k \equiv \left. \frac{\partial \mathbf{h}_k}{\partial \mathbf{w}} \right|_{\mathbf{w}=\mathbf{w}_k^f} . \tag{26}$$

Equation 25 is obtained by linearizing (24) about the forecast state and using (1) and (2). The accuracy of the approximation (25) depends on the size of the forecast errors; it is exact for linear observation operators.

The residual mean and covariance are easily obtained from (25):

$$\langle \mathbf{v}_k \rangle \approx \mathbf{b}_k^o - \mathbf{H}_k \mathbf{b}_k^f, \tag{27}$$

$$\left\langle (\mathbf{v}_k - \langle \mathbf{v}_k \rangle)(\mathbf{v}_k - \langle \mathbf{v}_k \rangle)^T \right\rangle \approx \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T. \tag{28}$$

We used the additional approximation $\langle \mathbf{H}_k \cdot \rangle \approx \mathbf{H}_k \langle \cdot \rangle$; for linear observation operators (27) and (28) are both exact. Compare (28) with (10); specification of optimal weights in the analysis update requires complete knowledge of the residual covariance.

3.2 A state-space description of forecast bias.

We now assume that there exists a subset of the observing system for which bias is negligible:

$$\mathbf{b}_k^o \approx 0, \tag{29}$$

or, rather, that $|\mathbf{b}_k^o| \ll |\mathbf{H}_k \mathbf{b}_k^f|$ in some meaningful sense. This amounts to the requirement that systematic errors, if any, have been effectively removed from the observations. In that case (25) can be re-written

$$\mathbf{v}_k = -\mathbf{H}_k \mathbf{b}_k^f + \tilde{\boldsymbol{\eta}}_k, \tag{30}$$

where $\tilde{\boldsymbol{\eta}}_k$ is a noise term whose first- and second-order statistics are

$$\langle \tilde{\boldsymbol{\eta}}_k \rangle \approx 0, \tag{31}$$

$$\langle \tilde{\boldsymbol{\eta}}_k \tilde{\boldsymbol{\eta}}_k^T \rangle \approx \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T. \tag{32}$$

This follows from (25) and (29) by noting that $\tilde{\boldsymbol{\eta}}_k = \mathbf{v}_k - \langle \mathbf{v}_k \rangle$ and using (28).

Equation (30) can be regarded as a *measurement model* for the forecast bias $\mathbf{b}_k^f$. It expresses the relationship between the observations, the forecast, and the actual forecast bias under the assumption (29). If observations alone are insufficient to completely determine forecast bias, they must be supplemented with additional information. We therefore introduce a *state model* for $\mathbf{b}_k^f$ which describes its evolution in time. Formulation of the state model in fact amounts to an explicit definition of the quantity we wish to estimate, i.e., of our notion of forecast bias.

Our practical goal is to estimate the time-mean forecast error, averaged over a time period which exceeds synoptic time scales. By definition, this quantity is approximately constant in time, so that a reasonable state model for $\mathbf{b}_k^f$ is *the persistence model*

$$\mathbf{b}_k^f = \mathbf{b}_{k-1}^f. \tag{33}$$

This model will serve to predict forecast bias for time t_k based on a previous bias estimate valid for time t_{k-1}.

Forecast errors are state-dependent, and the evolution in time of forecast bias is therefore likely to be more complex than the persistence model (33) suggests. For example, the presence of a systematic error in the convective parameterization of a forecast model will result in systematic but transient short-term forecast errors in convectively active regions. Tibaldi and Molteni (1990) and Miyakoda and Sirutis (1990) discuss systematic forecast errors which occur during the onset of blocking, and their impact on forecast skill. It will be a challenge to express this type of information explicitly in terms of a bias evolution model of a more general form, say,

$$\mathbf{b}_k^f = \mathbf{b}_{k-1}^f + \mathbf{g}(\mathbf{w}_{k-1}^t), \tag{33$'$}$$

where $\mathbf{g}$ is some nonlinear operator.

Equations (30) and (33) (or (33′)) together constitute a *state-space description* (Anderson and Moore 1979) of the forecast bias $\mathbf{b}_k^f$. Given such a description, the estimation of this quantity becomes a standard problem in estimation theory. Griffith and Nichols (1996) pursue a similar approach, but in the context of variational data assimilation. They propose to extend the *variational continuous assimilation method* (Derber 1989) by introducing a deterministic evolution model for model error, analogous to (33′). The model error is then treated as part of the control variable in the variational formulation of the data assimilation problem, and can be estimated along with the forecast trajectory using adjoint techniques.

3.3 Sequential estimation of forecast bias.

A sequential bias estimation algorithm producing estimates $\hat{\mathbf{b}}_k$ of the forecast bias $\mathbf{b}_k^f$ can be defined recursively as follows. Given a previous bias estimate $\hat{\mathbf{b}}_{k-1}$, the persistence model (33) predicts the forecast bias at time t_k simply by

$$\hat{\mathbf{b}}_k^- = \hat{\mathbf{b}}_{k-1}. \tag{34}$$

In case of the more general model (33′) the bias prediction might be

$$\hat{\mathbf{b}}_k^- = \hat{\mathbf{b}}_{k-1} + \mathbf{g}(\hat{\mathbf{b}}_{k-1}, \mathbf{w}_{k-1}^a). \tag{34$'$}$$

An updated estimate $\hat{\mathbf{b}}_k$ of forecast bias can be obtained by combining the bias prediction $\hat{\mathbf{b}}_k^-$ with the measurements provided by (30). It is easy to show from (30–32) that the least-variance unbiased linear combination of prediction and measurements is given by

$$\hat{\mathbf{b}}_k = \hat{\mathbf{b}}_k^- - \mathbf{L}_k \left[\mathbf{v}_k + \mathbf{H}_k \hat{\mathbf{b}}_k^-\right], \tag{35}$$

with

$$\mathbf{L}_k = \mathbf{P}_k^{b-} \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k\right]^{-1}, \tag{36}$$

where $\mathbf{P}_k^{b-}$ is the error covariance of the bias estimate $\hat{\mathbf{b}}_k^-$:

$$\mathbf{P}_k^{b-} \equiv \left\langle (\hat{\mathbf{b}}_k^- - \mathbf{b}_k^f)(\hat{\mathbf{b}}_k^- - \mathbf{b}_k^f)^T \right\rangle. \tag{37}$$

The algorithm must be initialized with an *a priori* bias estimate $\hat{\mathbf{b}}_0$, and it requires specification of the error covariances $\mathbf{P}_k^{b-}$.

For a linear stochastic-dynamic bias evolution model (in particular, for the persistence model (33)) it is possible to derive recursions for the covariances $\mathbf{P}_k^{b-}$, as we shall show in section 4. Supplemented by these recursions, the algorithm (34–36) is just the Kalman filter for the system (30, 33). It will be more practical, however, to specify the covariance $\mathbf{P}_k^{b-}$ directly—that is, without recourse to the covariance equations—analogous to the direct modeling of forecast and observation error covariances in operational data assimilation systems. We will return to the issue of estimation error covariance modeling in section 5.

In case of a linear bias model the bias estimate $\hat{\mathbf{b}}_k$ defined by (35) is unbiased, provided the observations are unbiased:

$$\langle \hat{\mathbf{b}}_k \rangle = \mathbf{b}_k^f. \tag{38}$$

This statement follows directly from (35) combined with (29), and does not depend on the particular gain $\mathbf{L}_k$. The least-variance property of the estimator, on the other hand, holds only if the error covariances $\mathbf{P}_k^{b-}$, $\mathbf{P}_k^f$, and $\mathbf{R}_k$ are correctly specified in (36). Actual implementations of the algorithm will generally be suboptimal.

Stability properties of the sequential bias estimation algorithm can be stated in terms of stability properties of the Kalman filter. For linear bias models the convergence of $\hat{\mathbf{b}}_k$ to $\mathbf{b}_k^f$ (in the statistical mean-square sense) depends on observability and controllability properties of the state-space system (30,33 or 33′). In practical terms, and for the persistence model (33) in particular, this means that the (unbiased) observing system must provide sufficient coverage during the maximum time interval over which forecast bias can be presumed constant. Bias estimates at locations where no unbiased observations are available will be determined partly by the *a priori* bias estimate $\hat{\mathbf{b}}_0$ there, and partly by the specification of the error covariances $\mathbf{P}_k^{b-}$ between locations within and without the observed regions.

4 Bias estimation theory

In this section we summarize the approach to bias estimation first developed by Friedland (1969) and subsequently clarified and extended by others (Tacker and Lee 1972; Mendel 1976; Friedland 1978; Ignagni 1981; Ignagni 1990; Zhou *et al.* 1993). The work of these authors provides a rigorous framework for the sequential bias estimation algorithm presented in the previous section, and can be applied more generally to the estimation of model error parameters and/or observational bias.

4.1 General framework.

Friedland (1969) considered the problem of estimating the true state $\mathbf{w}_k^t$ of a linear stochastic-dynamic process in the presence of a set of constant (but unknown) bias parameters $\boldsymbol{\beta}$. Other than being constant, the bias parameters are rather generally defined, and may affect both the state model and the measurement model. In particular, it is not assumed that observations are unbiased. The theory has been developed for both continuous and discrete processes; here we present only the latter.

The framework assumes linear stochastic-dynamic state and measurement models of the form:

$$\mathbf{w}_k^t = \mathbf{A}_k \mathbf{w}_{k-1}^t + \mathbf{B}_k \boldsymbol{\beta} + \widetilde{\boldsymbol{\xi}}_k, \tag{39}$$

$$\mathbf{w}_k^o = \mathbf{H}_k \mathbf{w}_{k-1}^t + \mathbf{C}_k \boldsymbol{\beta} + \widetilde{\boldsymbol{\eta}}_k. \tag{40}$$

Here $\mathbf{A}_k, \mathbf{B}_k, \mathbf{H}_k, \mathbf{C}_k$ are known matrices of appropriate dimensions, and $\widetilde{\boldsymbol{\xi}}_k, \widetilde{\boldsymbol{\eta}}_k$ are mutually independent white Gaussian vector processes with known first- and second-order statistics

$$\langle \tilde{\boldsymbol{\xi}}_k \rangle = 0, \qquad \left\langle \tilde{\boldsymbol{\xi}}_k \tilde{\boldsymbol{\xi}}_k^T \right\rangle = \mathbf{Q}_k, \tag{41}$$

$$\langle \tilde{\boldsymbol{\eta}}_k \rangle = 0, \qquad \left\langle \tilde{\boldsymbol{\eta}}_k \tilde{\boldsymbol{\eta}}_k^T \right\rangle = \mathbf{R}_k. \tag{42}$$

If the bias parameters $\boldsymbol{\beta}$ were known (or if $\mathbf{B}_k \equiv \mathbf{C}_k \equiv 0$) the optimal state estimate at time t_k based on all observations up to that time would be given by the usual Kalman filter equations.

Note that the bias parameters can enter the problem in different ways, depending on the definition of the matrices $\mathbf{B}_k$ and $\mathbf{C}_k$. When $\mathbf{B}_k \equiv 0$ the state model (39) is unbiased; in our application this corresponds to an unbiased forecast model. When $\mathbf{C}_k \equiv 0$ the observations are unbiased. Generally, the term $\mathbf{B}_k\boldsymbol{\beta}$ represents the effect of unknown model error parameters entering into the state evolution. The bias vector $\boldsymbol{\beta}$ may consist of just a few parameters—say, unknown spectral coefficients of model error—or it may be dimensionally compatible with the true state $\mathbf{w}_k^t$.

4.2 Optimal state estimation in the presence of bias.

Optimal estimates of the true state $\mathbf{w}_k^t$ and the bias parameters $\boldsymbol{\beta}$ can be obtained by applying the standard technique of augmenting the state vector with the bias parameters (e.g., Jazwinski 1970, section 8.4). Linear state and measurement models for the augmented state follow from (39) and (40) together with the statement that the bias parameters are constant. The Kalman filter for this system then simultaneously provides the optimal estimates of $\mathbf{w}_k^t$ and $\boldsymbol{\beta}$. The obvious drawback to this approach is that it is not a simple matter to modify an existing implementation of a state estimation algorithm by introducing state augmentation.

Friedland showed that the Kalman filter equations for the augmented state are algebraically equivalent to two loosely coupled sets of recursions, resulting in a two-stage estimation algorithm. The first stage consists of the usual filter equations for the state $\mathbf{w}_k^t$, obtained by ignoring the bias altogether. The second stage provides estimates of the bias parameters $\boldsymbol{\beta}$ based on the output of the first stage. Results from the two stages can then be combined to produce the optimal (unbiased) state estimates.

This two-stage approach to concurrent state and bias estimation is known as *separate-bias estimation* in the literature. The first stage in the algorithm was originally called the *bias-free state estimator* by Friedland, since none of the equations in this stage involve bias estimates. We have found this terminology to be potentially confusing since it suggests that bias-free state estimates are unbiased, which is not actually the case. We therefore prefer to use the term *bias-blind state estimator*, which more clearly indicates that bias is present yet ignored in that part of the algorithm.

Friedland's two-stage approach is attractive for many applications because the bias estimator can be implemented as a supplemental component to an existing (bias-blind) state estimator: the design of the state estimator is unaffected by the addition of the bias estimator. The latter can be activated as needed, e.g. when output diagnostics indicate significant bias problems. We include the complete set of algorithms and some important properties here without proof; see Friedland (1969; 1978) and Ignagni (1981; 1990) for details.

Bias-blind state estimator.

The bias-blind state estimates $\mathbf{w}_k^f, \mathbf{w}_k^a$ are given by

$$\mathbf{w}_k^f = \mathbf{A}_k \mathbf{w}_{k-1}^a, \tag{43}$$

$$\mathbf{w}_k^a = \mathbf{w}_k^f + \mathbf{K}_k \left[\mathbf{w}_k^o - \mathbf{H}_k \mathbf{w}_k^f \right], \tag{44}$$

where the gain $\mathbf{K}_k$ is

$$\mathbf{K}_k = \mathbf{S}_k^f \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{S}_k^f \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1}, \tag{45}$$

and $\mathbf{S}_k^f$ is defined recursively by

$$\mathbf{S}_k^f = \mathbf{A}_k \mathbf{S}_{k-1}^a \mathbf{A}_k^T + \mathbf{Q}_k, \tag{46}$$

$$\mathbf{S}_k^a = \left[\mathbf{I} - \mathbf{K}_k \mathbf{H}_k \right] \mathbf{S}_k^f. \tag{47}$$

In the absence of bias (43–47) are just the Kalman filter equations.

Bias estimator.

Bias parameter estimates $\hat{\boldsymbol{\beta}}_k$ are given by

$$\hat{\boldsymbol{\beta}}_k = \hat{\boldsymbol{\beta}}_{k-1} + \mathbf{K}_k^\beta \left[\mathbf{w}_k^o - \mathbf{H}_k \mathbf{w}_k^f - \mathbf{T}_k \hat{\boldsymbol{\beta}}_{k-1} \right], \tag{48}$$

where the gain $\mathbf{K}_k^\beta$ is

$$\mathbf{K}_k^\beta = \mathbf{P}_{k-1}^\beta \mathbf{T}_k^T \left[\mathbf{T}_k \mathbf{P}_{k-1}^\beta \mathbf{T}_k^T + \mathbf{H}_k \mathbf{S}_k^f \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1}, \tag{49}$$

and $\mathbf{P}_k^\beta$ is defined recursively by

$$\mathbf{P}_k^\beta = \left[\mathbf{I} - \mathbf{K}_k^\beta \mathbf{T}_k \right] \mathbf{P}_{k-1}^\beta. \tag{50}$$

The matrix $\mathbf{T}_k$ is defined by the following set of recursions:

$$\mathbf{T}_k = \mathbf{H}_k \mathbf{U}_k + \mathbf{C}_k, \tag{51}$$

$$\mathbf{U}_k = \mathbf{A}_k \mathbf{V}_{k-1} + \mathbf{B}_k, \tag{52}$$

$$\mathbf{V}_k = \mathbf{U}_k - \mathbf{K}_k \mathbf{T}_k. \tag{53}$$

The last equation in this set depends on the state estimator gain $\mathbf{K}_k$.

Bias correction.

Unbiased state estimates $\tilde{\mathbf{w}}_k^f, \tilde{\mathbf{w}}_k^a$ are obtained by

$$\tilde{\mathbf{w}}_k^f = \mathbf{w}_k^f + \mathbf{U}_k\hat{\boldsymbol{\beta}}_{k-1}, \tag{54}$$

$$\tilde{\mathbf{w}}_k^a = \mathbf{w}_k^a + \mathbf{V}_k\hat{\boldsymbol{\beta}}_k. \tag{55}$$

Initialization.

The recursions for the two estimators are initialized by specifying *a priori* state and bias estimates $\mathbf{w}_0^a$ and $\hat{\boldsymbol{\beta}}_0$ as well as the matrices $\mathbf{S}_0^a$, $\mathbf{P}_0^\beta$, and $\mathbf{V}_0$. Generally, if it is assumed that

$$\mathbf{w}_0^t = \mathbf{w}_0 + \mathbf{M}_0\boldsymbol{\beta}, \qquad \left\langle \mathbf{w}_0\boldsymbol{\beta}^T \right\rangle = 0 \tag{56}$$

with $\mathbf{M}_0$ a known matrix, then

$$\mathbf{S}_0^a = \left\langle (\mathbf{w}_0^a - \mathbf{w}_0)(\mathbf{w}_0^a - \mathbf{w}_0)^T \right\rangle, \tag{57}$$

$$\mathbf{P}_0^\beta = \left\langle (\hat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_0 - \boldsymbol{\beta})^T \right\rangle, \tag{58}$$

$$\mathbf{V}_0 = \mathbf{M}_0. \tag{59}$$

4.3 Properties and some extensions.

Friedland (1969) showed that the $\hat{\boldsymbol{\beta}}_k$ defined by (48–53) are optimal estimates of the bias parameters $\boldsymbol{\beta}$, given all observations up to and including time t_k. The matrices $\mathbf{P}_k^\beta$ are the actual estimation error covariances:

$$\mathbf{P}_k^\beta = \left\langle (\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta})^T \right\rangle. \tag{60}$$

The unbiased state estimates $\tilde{\mathbf{w}}_k^f, \tilde{\mathbf{w}}_k^a$ defined by (54, 55) are optimal, with error covariances given by

$$\mathbf{P}_k^f \equiv \left\langle (\tilde{\mathbf{w}}_k^f - \mathbf{w}_k^t)(\tilde{\mathbf{w}}_k^f - \mathbf{w}_k^t)^T \right\rangle = \mathbf{S}_k^f + \mathbf{U}_k \mathbf{P}_{k-1}^\beta \mathbf{U}_k^T, \tag{61}$$

$$\mathbf{P}_k^a \equiv \left\langle (\tilde{\mathbf{w}}_k^a - \mathbf{w}_k^t)(\tilde{\mathbf{w}}_k^a - \mathbf{w}_k^t)^T \right\rangle = \mathbf{S}_k^a + \mathbf{V}_k \mathbf{P}_k^\beta \mathbf{V}_k^T. \tag{62}$$

The cross-covariances between the state estimates and the bias estimates are

$$\left\langle (\tilde{\mathbf{w}}_k^f - \mathbf{w}_k^t)(\hat{\boldsymbol{\beta}}_{k-1} - \boldsymbol{\beta})^T \right\rangle = \mathbf{U}_k \mathbf{P}_{k-1}^\beta, \tag{63}$$

$$\left\langle (\tilde{\mathbf{w}}_k^a - \mathbf{w}_k^t)(\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta})^T \right\rangle = \mathbf{V}_k \mathbf{P}_k^\beta. \tag{64}$$

Unless $\boldsymbol{\beta} = 0$ the bias-blind state estimator (43–47) is suboptimal and produces biased state estimates $\mathbf{w}_k^f, \mathbf{w}_k^a$. The quantities $\mathbf{S}_k^f, \mathbf{S}_k^a$ are covariance matrices, i.e., they are in fact positive semidefinite, but (61, 62) show that they differ from the actual estimation error covariances $\mathbf{P}_k^f, \mathbf{P}_k^a$ for the unbiased estimates $\tilde{\mathbf{w}}_k^f, \tilde{\mathbf{w}}_k^a$. One can prove, however, that the covariances $\mathbf{S}_k^f, \mathbf{S}_k^a$ do converge to $\mathbf{P}_k^f, \mathbf{P}_k^a$ as $k \to \infty$, provided the model system (39, 40) is uniformly completely observable and uniformly completely controllable. In that case the second term in each of the equations (61, 62) approaches zero as the bias parameter estimates converge, since the $\mathbf{U}_k, \mathbf{V}_k$ defined by (51–53) will remain bounded due to the stability of the bias-blind state estimator.

Several extensions to the two-stage state and bias estimation algorithm have been proposed based on natural generalizations of the linear, constant-bias framework provided by the model system (39, 40). Various authors (Tacker and Lee 1972; Ignani 1990) have considered model systems in which the bias parameters $\boldsymbol{\beta}$ are allowed to vary in time. Optimal bias and state estimates can still be obtained using simple extensions of Friedland's two-stage estimator, provided the variation in time of the bias parameters can be modeled and the resulting model system is still linear. Others (Mendel 1976; Zhou *et al.* 1993) have proposed modifications to the algorithms designed to handle nonlinear state models; the resulting estimates are, of course, suboptimal.

4.4 Forecast bias estimation.

Using our definition (3) of forecast bias in section 2,

$$\mathbf{b}_k^f \equiv \langle \mathbf{w}_k^f - \mathbf{w}_k^t \rangle \tag{65}$$

$$= \langle \tilde{\mathbf{w}}_k^f - \mathbf{w}_k^t \rangle - \langle \mathbf{U}_k \hat{\boldsymbol{\beta}}_{k-1} \rangle \tag{66}$$

$$= -\mathbf{U}_k \boldsymbol{\beta} \tag{67}$$

where we used (54) and the fact that $\tilde{\mathbf{w}}_k^f$ and $\hat{\boldsymbol{\beta}}_{k-1}$ are unbiased estimates of the true state $\mathbf{w}_k^t$ and the bias parameters $\boldsymbol{\beta}$, respectively. In fact, $\hat{\boldsymbol{\beta}}_k$ is an unbiased estimate of $\boldsymbol{\beta}$ as well, so that

$$\hat{\mathbf{b}}_k^- \equiv -\mathbf{U}_k \hat{\boldsymbol{\beta}}_{k-1}, \tag{68}$$

$$\hat{\mathbf{b}}_k \equiv -\mathbf{U}_k \hat{\boldsymbol{\beta}}_k \tag{69}$$

are both estimates of forecast bias $\mathbf{b}_k^f$. Their error covariances are

$$\mathbf{P}_k^{b^-} \equiv \left\langle (\hat{\mathbf{b}}_k^- - \mathbf{b}_k^f)(\hat{\mathbf{b}}_k^- - \mathbf{b}_k^f)^T \right\rangle = \mathbf{U}_k \mathbf{P}_{k-1}^{\beta} \mathbf{U}_k^T, \tag{70}$$

$$\mathbf{P}_k^{b} \equiv \left\langle (\hat{\mathbf{b}}_k - \mathbf{b}_k^f)(\hat{\mathbf{b}}_k - \mathbf{b}_k^f)^T \right\rangle = \mathbf{U}_k \mathbf{P}_k^{\beta} \mathbf{U}_k^T, \tag{71}$$

using (60).

The estimate $\hat{\mathbf{b}}_k^-$ *predicts* forecast bias at time t_k based on data prior to t_k: the corrected (unbiased) forecast (54) is

$$\tilde{\mathbf{w}}_k^f = \mathbf{w}_k^f - \hat{\mathbf{b}}_k^-. \tag{72}$$

The second estimate $\hat{\mathbf{b}}_k$ is an update of the first, using additional data available at time t_k. Optimality of the bias estimator implies that these are the least-variance estimates given the observations upon which they are based; in particular, the updated estimate $\hat{\mathbf{b}}_k$ is more accurate than $\hat{\mathbf{b}}_k^-$.

We will now review the sequential forecast bias estimation algorithm (34–36) developed in section 3, using the theory presented here. Consider the special case in which the measurement model (40) is unaffected by the bias parameters, i.e.,

$$\mathbf{C}_k \equiv 0. \tag{73}$$

This corresponds to the assumption that observations are unbiased. Equations (51–53) then reduce to

$$\mathbf{T}_k = \mathbf{H}_k \mathbf{U}_k, \tag{74}$$

$$\mathbf{U}_k = \mathbf{A}_k \left[\mathbf{I} - \mathbf{K}_{k-1}\mathbf{H}_{k-1}\right] \mathbf{U}_{k-1} + \mathbf{B}_k. \tag{75}$$

Multiplying (48) by $-\mathbf{U}_k$ and using (68–70) and (74), we obtain the forecast bias update equation

$$\hat{\mathbf{b}}_k = \hat{\mathbf{b}}_k^- - \mathbf{L}_k \left[\mathbf{w}_k^o - \mathbf{H}_k \mathbf{w}_k^f + \mathbf{H}_k \hat{\mathbf{b}}_k^-\right], \tag{76}$$

with

$$\mathbf{L}_k \equiv \mathbf{U}_k \mathbf{K}_k^\beta \tag{77}$$

$$= \mathbf{P}_k^{b-} \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{P}_k^{b-} \mathbf{H}_k^T + \mathbf{H}_k \mathbf{S}_k^f \mathbf{H}_k^T + \mathbf{R}_k\right]^{-1}. \tag{78}$$

Compare (76, 78) with (35, 36); they are identical by virtue of (24), (61), and (70). The covariance update equation implied by (50) is

$$\mathbf{P}_k^b = \left[\mathbf{I} - \mathbf{L}_k \mathbf{H}_k\right] \mathbf{P}_k^{b-}, \tag{79}$$

where we used (77) and (70, 71).

Equations for the forecast bias prediction $\hat{\mathbf{b}}_k^-$ and its error covariance $\mathbf{P}_k^{b-}$ can be similarly obtained. For prediction, however, it is necessary to define the precise relationship between the bias parameters $\boldsymbol{\beta}$ and the evolution of the true state, represented by the matrix $\mathbf{B}_k$ in the state model (39). In section 3 we took a somewhat different approach by directly modeling the evolution in time of the forecast bias itself. In case of the persistence model (33) this is easily reconciled with the present theory, as follows. If the forecast bias $\mathbf{b}_k^f$ is presumed constant in time, (67) implies that

$$\mathbf{U}_k \equiv \mathbf{U} = \text{const.} \tag{80}$$

But then (68, 69) together imply

$$\hat{\mathbf{b}}_k^- = \hat{\mathbf{b}}_{k-1}, \tag{81}$$

which is just (34). Similarly, (70, 71) together imply

$$\mathbf{P}_k^{b-} = \mathbf{P}_{k-1}^b. \tag{82}$$

Note from (75) that (80) corresponds to the particular choice

$$\mathbf{B}_k = \left\{\mathbf{I} - \mathbf{A}_k\left[\mathbf{I} - \mathbf{K}_{k-1}\mathbf{H}_{k-1}\right]\right\}\mathbf{U} \tag{83}$$

in the state model (39), for an arbitrary (but constant) matrix $\mathbf{U}$. It would be farfetched to assume that the model (39) with this choice of $\mathbf{B}_k$ provides a realistic description of the actual state evolution. For this and other practical reasons it makes sense to bypass the covariance equations (79, 82) altogether and instead approximate the estimation error covariances directly, as is usually done in operational data assimilation systems.

4.5 The unbiased analysis equation revisited.

Still assuming unbiased observations ($\mathbf{C}_k \equiv 0$) it follows from (51) and (53) that

$$\mathbf{V}_k = [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\,\mathbf{U}_k \tag{84}$$

which, together with (55, 69) implies

$$\widetilde{\mathbf{w}}_k^a = \mathbf{w}_k^a - [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\,\hat{\mathbf{b}}_k, \tag{85}$$

showing how the updated forecast bias estimate $\hat{\mathbf{b}}_k$ may be used to obtain the optimal (unbiased) analysis.

Equation (85) can be used to re-derive the unbiased analysis equation (8), as follows. Using (44) and (76),

$$\widetilde{\mathbf{w}}_k^a = \mathbf{w}_k^f + \mathbf{K}_k\left[\mathbf{w}_k^o - \mathbf{H}_k\mathbf{w}_k^f\right] - [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\,\hat{\mathbf{b}}_k \tag{86}$$

$$= \widetilde{\mathbf{w}}_k^f + \{\mathbf{K}_k + [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\,\mathbf{L}_k\}\left[\mathbf{w}_k^o - \mathbf{H}_k\widetilde{\mathbf{w}}_k^f\right]. \tag{87}$$

The covariance relation (61) can also be written as

$$\mathbf{P}_k^f = \mathbf{S}_k^f + \mathbf{P}_k^{b-}, \tag{88}$$

and this together with (49), (78) implies

$$\{\mathbf{K}_k + [\mathbf{I} - \mathbf{K}_k\mathbf{H}_k]\,\mathbf{L}_k\} = \mathbf{P}_k^f\mathbf{H}_k^T\left[\mathbf{H}_k\mathbf{P}_k^f\mathbf{H}_k^T + \mathbf{R}_k\right]^{-1}, \tag{89}$$

which proves (8), since $\widetilde{\mathbf{w}}_k^o = \mathbf{w}_k^o$ here (observations are unbiased).

Another useful way to express the unbiased analysis equation is in terms of the quantity

$$\widetilde{\widetilde{\mathbf{w}}}_k^f \equiv \mathbf{w}_k^f - \hat{\mathbf{b}}_k \tag{90}$$

which is the *a posteriori* unbiased forecast. From (86) we simply obtain

$$\widetilde{\mathbf{w}}_k^a = \widetilde{\widetilde{\mathbf{w}}}_k^f + \mathbf{K}_k \left[\mathbf{w}_k^o - \mathbf{H}_k \widetilde{\widetilde{\mathbf{w}}}_k^f \right]. \tag{91}$$

5 Implementation in practice

In this section we summarize the on-line and off-line bias estimation algorithms and address some practical issues. We also present the results of a simple numerical experiment performed with a linear one-dimensional shallow-water model.

5.1 Selection of observations for bias estimation.

The basic assumption leading to the ability to estimate forecast bias is that unbiased observations are available for this purpose. In practical terms this means that observational bias must have been effectively removed, so that any remaining systematic component of the time-averaged observed-minus-forecast residuals can be attributed to forecast bias. This requirement is not realistic for all data types used in operational data assimilation systems. In practice one may therefore need to define a limited subset of observations to be used for the purpose of estimating forecast bias.

We will use the superscript notation $^{|}$ to indicate that a quantity is associated with this restricted subset. In particular, $\mathbf{w}_k^{o|}$ denotes the vector of unbiased observations at time t_k; it is a subset of the full set of observations $\mathbf{w}_k^o$. Similarly let $\mathbf{h}_k^|, \mathbf{H}_k^|$, and $\mathbf{R}_k^|$ be the corresponding restrictions of the observation operator $\mathbf{h}_k$, the linearized observation operator $\mathbf{H}_k$, and the observation error covariance $\mathbf{R}_k$, respectively.

5.2 Summary of the algorithms.

The off-line algorithm produces estimates of forecast bias based on stored output from an assimilation system. The on-line algorithm, on the other hand, utilizes current forecast bias predictions and updates in order to produce unbiased forecasts and analyses. Bias estimates are based exclusively on observations which are considered unbiased. The computational cost of bias estimation is roughly that of a single solution of the analysis equation for a limited number of observations.

Off-line forecast bias estimation.

bias prediction:

$$\hat{\mathbf{b}}_k^- = \hat{\mathbf{b}}_{k-1} + \mathbf{g}(\hat{\mathbf{b}}_{k-1}, \mathbf{w}_{k-1}^a) \tag{92}$$

$$\tilde{\mathbf{w}}_k^f = \mathbf{w}_k^f - \hat{\mathbf{b}}_k^- \tag{93}$$

bias update:

$$\hat{\mathbf{b}}_k = \hat{\mathbf{b}}_k^- - \mathbf{L}_k \left[\mathbf{w}_k^{o|} - \mathbf{h}_k^|(\tilde{\mathbf{w}}_k^f)\right] \tag{94}$$

$$\mathbf{L}_k = \mathbf{S}_k^{b-} \mathbf{H}_k^{|\,T} \left[\mathbf{H}_k^| \mathbf{S}_k^{b-} \mathbf{H}_k^{|\,T} + \mathbf{H}_k^| \mathbf{S}_k^f \mathbf{H}_k^{|\,T} + \mathbf{R}_k^|\right]^{-1} \tag{95}$$

On-line forecast bias estimation and correction.

forecast and bias prediction:

$$\mathbf{w}_k^f = \mathbf{A}_k \tilde{\mathbf{w}}_{k-1}^a \tag{96}$$

$$\hat{\mathbf{b}}_k^- = \hat{\mathbf{b}}_{k-1} + \mathbf{g}(\hat{\mathbf{b}}_{k-1}, \tilde{\mathbf{w}}_{k-1}^a) \tag{97}$$

$$\tilde{\mathbf{w}}_k^f = \mathbf{w}_k^f - \hat{\mathbf{b}}_k^- \tag{98}$$

bias update:

$$\hat{\mathbf{b}}_k = \hat{\mathbf{b}}_k^- - \mathbf{L}_k \left[\mathbf{w}_k^{o|} - \mathbf{h}_k^|(\tilde{\mathbf{w}}_k^f)\right] \tag{99}$$

$$\mathbf{L}_k = \mathbf{S}_k^{b-} \mathbf{H}_k^{|\,T} \left[\mathbf{H}_k^| \mathbf{S}_k^{b-} \mathbf{H}_k^{|\,T} + \mathbf{H}_k^| \mathbf{S}_k^f \mathbf{H}_k^{|\,T} + \mathbf{R}_k^|\right]^{-1} \tag{100}$$

$$\tilde{\tilde{\mathbf{w}}}_k^f = \mathbf{w}_k^f - \hat{\mathbf{b}}_k \tag{101}$$

analysis:

$$\tilde{\mathbf{w}}_k^a = \tilde{\tilde{\mathbf{w}}}_k^f + \mathbf{K}_k \left[\mathbf{w}_k^o - \mathbf{h}_k(\tilde{\tilde{\mathbf{w}}}_k^f)\right] \tag{102}$$

$$\mathbf{K}_k = \mathbf{S}_k^f \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{S}_k^f \mathbf{H}_k^T + \mathbf{R}_k\right]^{-1} \tag{103}$$

5.3 Covariance modeling.

According to the theory presented in section 4, the on-line estimates are optimal when

$$\mathbf{S}_k^f = \mathbf{P}_k^f - \mathbf{P}_k^{b-}, \tag{104}$$

$$\mathbf{S}_k^{b-} = \mathbf{P}_k^{b-}, \tag{105}$$

where

$$\mathbf{P}_k^f = \left\langle (\tilde{\mathbf{w}}_k^f - \mathbf{w}_k^t)(\tilde{\mathbf{w}}_k^f - \mathbf{w}_k^t)^T \right\rangle, \tag{106}$$

$$\mathbf{P}_k^{b-} = \left\langle (\hat{\mathbf{b}}_k^- - \mathbf{b}_k^f)(\hat{\mathbf{b}}_k^- - \mathbf{b}_k^f)^T \right\rangle. \tag{107}$$

Note that $\mathbf{S}_k^f$ *would* be the forecast error covariance in the optimal case if there were no bias, or if the bias were precisely known and properly handled in the analysis. The covariance $\mathbf{P}_k^f$, on the other hand, is the optimal forecast error covariance in case of a biased forecast model, and it is bounded from below by $\mathbf{S}_k^f$: it is clearly not possible to do as well when forecast bias is not precisely known.

In operational data assimilation systems forecast error covariances are approximated on the basis of various simplifying assumptions about the forecast error fields. For example, it is usually assumed that height error correlations at fixed pressure levels are isotropic, and that extra-tropical wind error fields are in geostrophic balance with the height errors. Remaining free parameters in the covariance formulations (such as error standard deviations and spatial de-correlation length scales) can then be estimated from past observed-minus-forecast residuals, although on-line estimation of some parameters based on current data is possible as well (Dee 1995).

A similar direct modeling approach could be used to approximate the forecast bias prediction error covariance $\mathbf{P}_k^{b-}$. Use of the Kalman filter recursions for these covariances in a practical application is not sensible, especially when forecast bias is being predicted by means of a simple heuristic model. The assumptions underlying covariance models for bias prediction errors $\hat{\mathbf{b}}_k^- - \mathbf{b}_k^f$ and for the random forecast errors $\tilde{\mathbf{w}}_k^f - \mathbf{w}_k^t$ should be consistent. For example, it is natural to assume that the systematic component of forecast error is approximately geostrophic as well in the middle latitudes.

Parameters for approximating the forecast error covariance $\mathbf{P}_k^f$ should be estimated from mean-zero observed-minus-forecast residuals

$$\tilde{\mathbf{v}}_k \equiv \mathbf{w}_k^o - \mathbf{h}_k(\tilde{\mathbf{w}}_k^f), \tag{108}$$

since

$$\left\langle \tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k^T \right\rangle \approx \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T. \tag{109}$$

A simple prescription for the bias prediction error covariance in terms of the forecast error covariance $\mathbf{P}_k^f$ might be

$$\mathbf{S}_k^{b-} = \gamma \mathbf{P}_k^f, \tag{110}$$

with $0 < \gamma < 1$ a free parameter. Consistent with (104) one should then take

$$\mathbf{S}_k^f = (1-\gamma)\mathbf{P}_k^f. \tag{111}$$

With this covariance model the gains $\mathbf{K}_k$ and $\mathbf{L}_k$ are given in terms of $\mathbf{P}_k^f$ and $\mathbf{R}_k$ by

$$\mathbf{K}_k = (1-\gamma)\mathbf{P}_k^f \mathbf{H}_k^T \left[(1-\gamma)\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1}, \tag{112}$$

$$\mathbf{L}_k = \gamma \mathbf{P}_k^f \mathbf{H}_k^{|\,T} \left[\mathbf{H}_k^{|} \mathbf{P}_k^f \mathbf{H}_k^{|\,T} + \mathbf{R}_k^{|} \right]^{-1}. \tag{113}$$

It should be relatively simple to use existing software for solving the analysis equations with these modified gains.

The parameter γ in our covariance model controls the stability of the bias estimates: the relative size of the bias updates is proportional to this parameter. An alternative interpretation is that γ determines the extent to which observational information is applied toward estimating the systematic rather than the random component of error. In the limit when $\gamma \to 0$ there are no bias updates and the on-line algorithm reduces to the usual analysis equation with a fixed (possibly zero) forecast bias correction. On the other hand, when $\gamma \to 1$ the observations are used exclusively for estimating forecast bias; the data assimilation scheme will then rely completely on the bias-corrected forecast. This would be appropriate if the forecast error were in fact entirely deterministic. Forecast error in practice consists of both systematic and random components; under some circumstances their relative magnitudes may be predictable and this information could be used to specify the parameter γ.

5.4 A numerical experiment.

We conclude this section with a simple application of the algorithm in a data assimilation system which is based on a discrete version of the linear one-dimensional shallow-water model

$$\frac{\partial h}{\partial t} + U\frac{\partial h}{\partial x} + H\frac{\partial u}{\partial x} - \frac{fU}{g}v = -\frac{1}{\tau}h + U\frac{\partial h_s}{\partial x}, \tag{114}$$

$$\frac{\partial u}{\partial t} + U\frac{\partial u}{\partial x} - fv + g\frac{\partial h}{\partial x} = -\frac{1}{\tau}u, \tag{115}$$

$$\frac{\partial v}{\partial t} + U\frac{\partial v}{\partial x} + fu = -\frac{1}{\tau}v. \tag{116}$$

Equations (114–116) are obtained by linearizing the nonlinear two-dimensional shallow-water equations at a fixed latitude φ about a geostrophic basic state, assuming periodic height- and wind-perturbations h, u, v depending on the longitudinal coordinate $0 \leq x \leq L$ only. The terms on the right-hand side represent linear damping and topographic forcing. Model parameters for our experiments are chosen to roughly represent quasi-geostrophic $500mb$ scales: $\varphi = 45°$ (model latitude), $L = 2\pi a \sin\varphi$ (length of the circle of latitude), $a = 6.371 \times 10^6 m$ (earth's radius), $g = 9.81 m^2 s^{-2}$ (gravity constant), $f = 2\Omega \sin\theta$ (Coriolis parameter), $\Omega = 2\pi/\text{day}$ (earth's rotation rate), $U = 17 ms^{-1}$ (mean zonal wind), $H = 8000m$ (mean height), and $\tau = 5$ days (damping time scale). The topography h_s is a smoothed version of the actual topography at 45N. We solve (114–116) numerically on a uniformly spaced grid with $M = 64$ points by means of the Richtmyer two-step formulation of the Lax-Wendroff scheme (Richtmyer and Morton 1967), using a time step of about 12 minutes. For quasi-geostrophic solutions of (114–116) the perturbation velocity component u is at least an order of magnitude smaller than v, and we will therefore completely ignore u in the presentation of our results.

The system (114–116) admits a stationary solution which is a function of the topography h_s. This stationary solution represents the climate of the true atmosphere in our experiment. The forecast model is defined by changing the specification of the topography (by using data at 48N instead), resulting in a fairly large climate error. Figure 2 shows the true topography (dotted curve) and forecast model topography (solid curve) in the top panel, and the $h-$ and $v-$components of the corresponding climates in the center and bottom panels, respectively. These climates were obtained by integrating (114–116), starting from a state of rest, for a period of 30 days. The systematic error in the forecast model which is due to the misspecification of topographic forcing will result in biased forecasts; the rms systematic height error in a 12-hour forecast is roughly 15m in this setup.

A single realization of the true atmospheric state evolution is simulated by adding

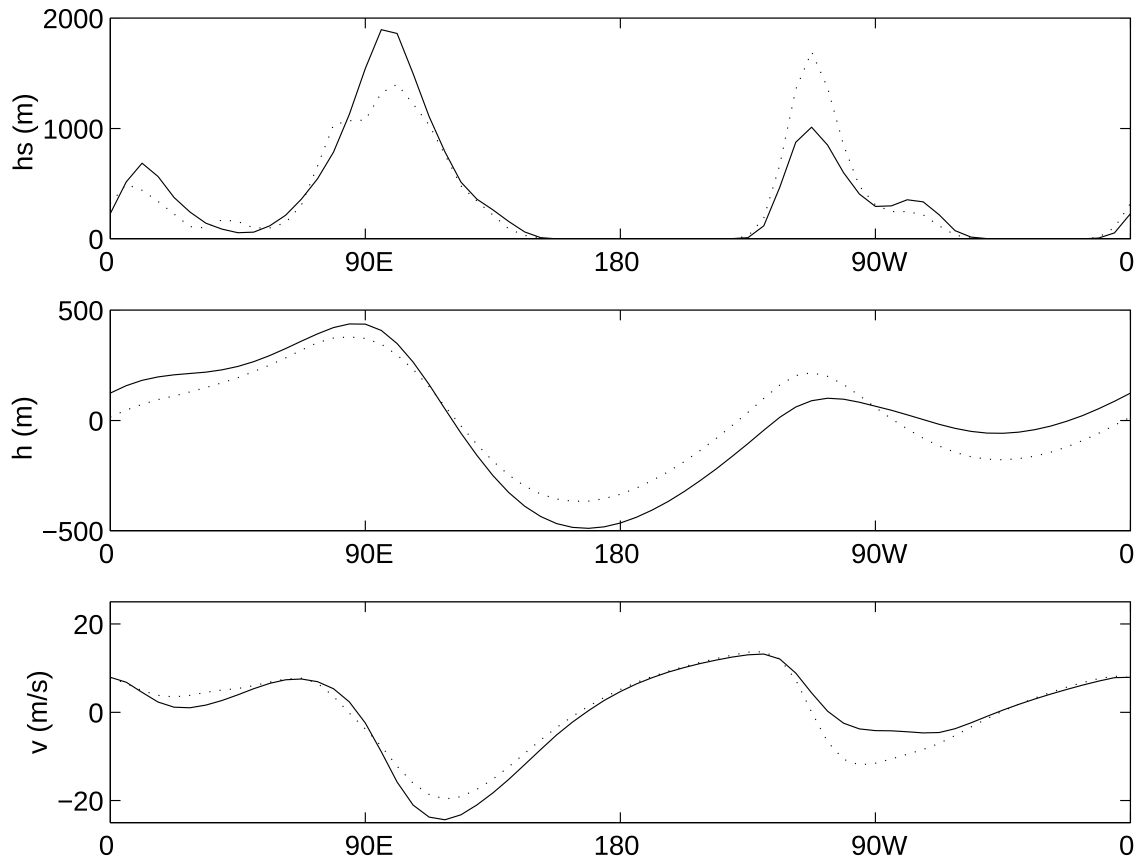

Figure 2: Top panel: true topography (dotted curve) and forecast model topography (solid curve). Center panel: $h-$component of true climate (dotted curve) and forecast model climate (solid curve). Bottom panel: $v-$component of true climate (dotted curve) and forecast model climate (solid curve).

a small, random forcing to the equations at each model time step. The forcing consists of a random linear combination of the slow modes of the system (114–116), the amplitude of the mode with wavenumber ω being normally distributed about zero with standard deviation $\sigma \exp[-(\omega/5)^2]$. The parameter σ is chosen such that the typical magnitude of the random component of the 12-hour forecast height error is comparable to that of the systematic component of error, i.e., about 15m. The simulated observing network provides 32 equally spaced height observations (at every other grid location) and 20 observations of the meridional wind component v (at every other grid location where $h_s > 0$) every 12 hours of simulation time; u is not observed. The height- and wind-observation error standard deviations specified in the data assimilation scheme are $40m$ and $4ms^{-1}$, respectively, although for our single-realization experiment we actually generate perfect observations.

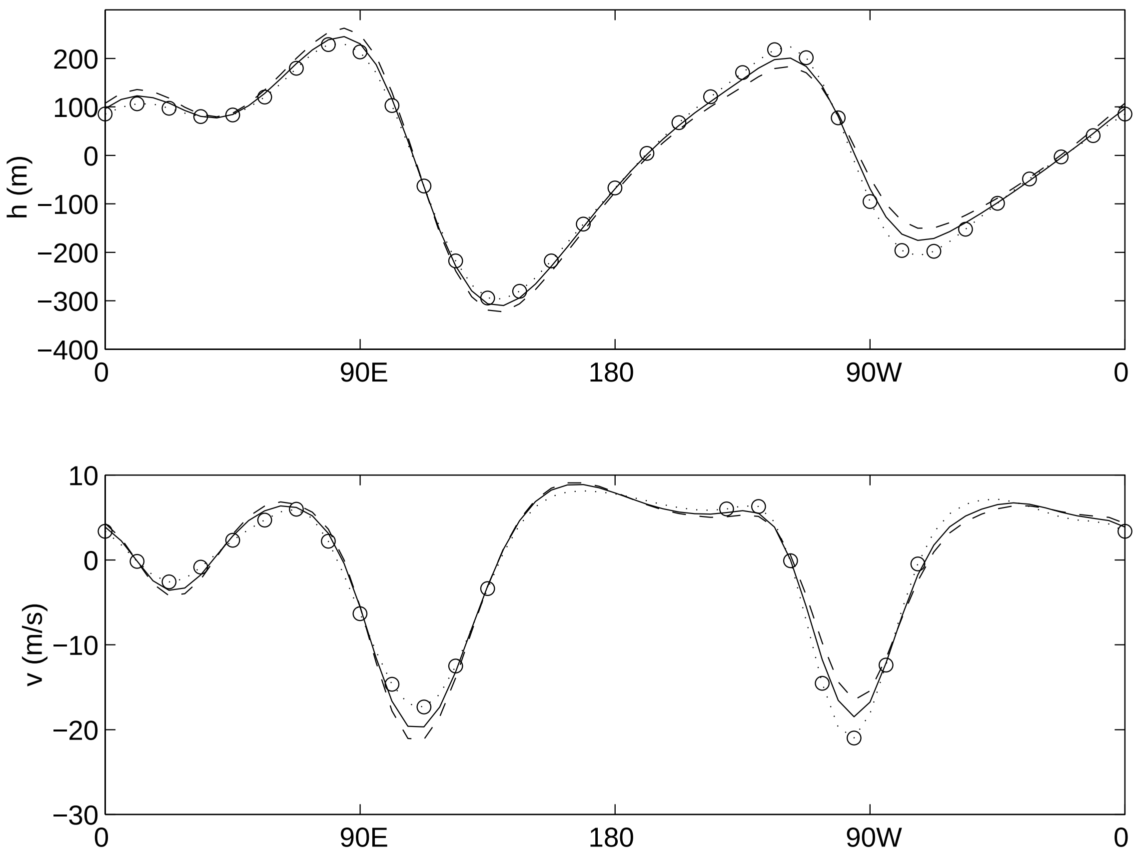

Figure 3: True state (dotted curve), bias-blind forecast (dashed curve), and bias-blind analysis (solid curve) after 48 hours of data assimilation. The $h-$component is shown in the top panel; the $v-$component in the bottom panel. Observations are taken every 12 hours as marked.

Rather than computing error covariances exactly in accordance with the theory, they will be prescribed as in (110) and (111) with $\gamma = 0.5$. The forecast error covariance $\mathbf{P}_k^f$

is modeled by assuming random isotropic height errors and geostrophically balanced wind errors. Geostrophy in (114–116) is expressed by $u = 0$, $v = \frac{g}{f}\frac{\partial h}{\partial x}$. Thus, our model for $\mathbf{P}_k^f$ is completely determined by two parameters: the height error standard deviation σ_h^f (taken to be constant in time and space) and a length scale parameter l for the spatial height error correlations (modeled by a simple power law). No attempt is made to tune or optimize these covariance parameters; we take $\sigma_h^f = 20m$ (roughly the rms norm of the 12-hour forecast height error) and $l = 1500km$ (roughly three grid cells).

Figure 3 shows a typical snapshot of the bias-blind state estimates, obtained by ignoring the forecast bias altogether. Heights are shown in the top panel; winds ($v-$component only) in the bottom panel. The solid curves represent the analyses produced at 48h. The dashed curves show the forecast based on the previous (36h) analysis, the dotted curves indicate the true state, and height- and wind-observations are marked. The forecast bias is visible in these plots, particularly near the troughs located to the east of the main topographic features. Note that the impact of the observations is positive, in the sense that the analysis errors are significantly smaller than the forecast errors.

Estimates of forecast bias were produced simultaneously with the bias-blind state estimates, based on all observations. In figure 4 we show the bias prediction (dashed curve) and the updated bias estimate (solid curve) after 48 hours of data assimilation, together with the true forecast error (dotted curve). The forecast error contains a random as well as a systematic component; only the latter is represented by the bias estimates. The bias-corrected state estimates are shown in figure 5. The analysis increments are now much smaller than those for the bias-blind estimates, since the

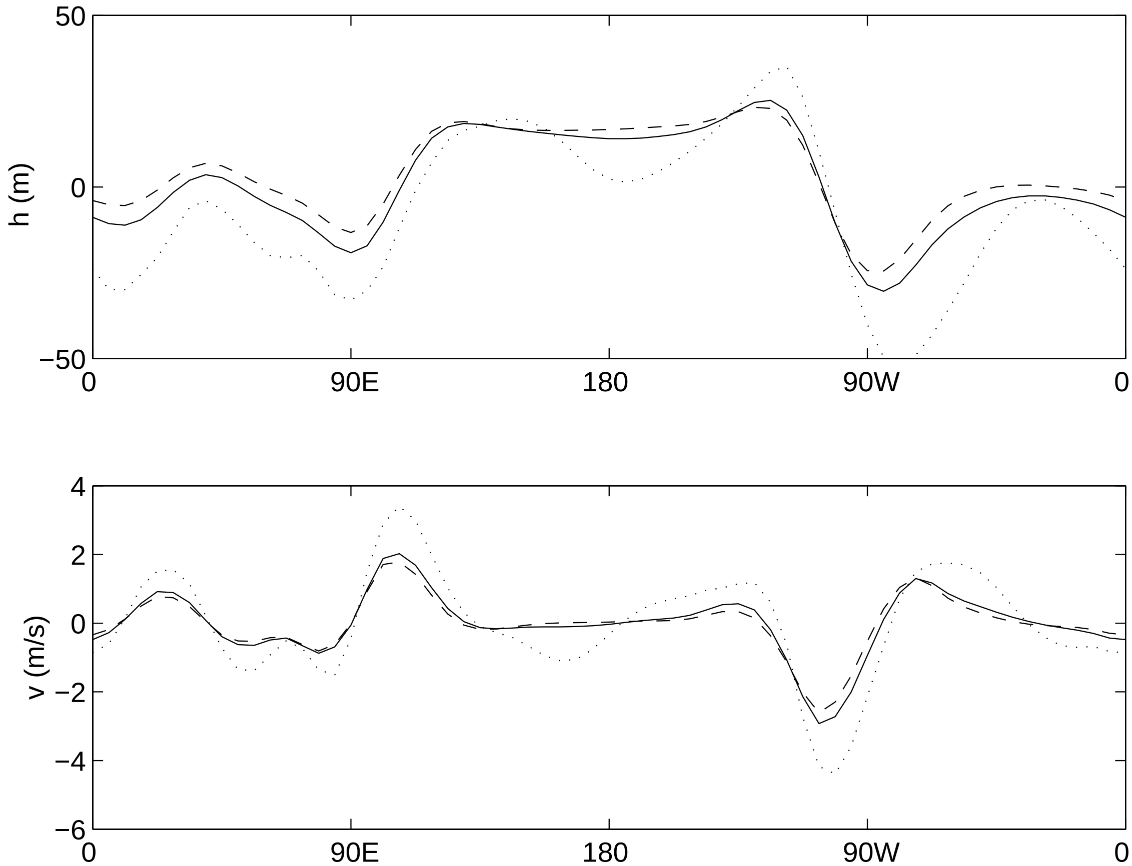

Figure 4: True forecast error (dotted curve), bias prediction (dashed curve), and updated bias estimate (solid curve) after 48 hours of data assimilation. The $h-$component is shown in the top panel; the $v-$component in the bottom panel.

forecast bias has been greatly reduced and only the random component of error is being corrected in the analysis.

Figure 6 shows the rms error evolution in time for the bias-blind (dashed curves) and for the bias-corrected (solid curves) state estimates, during the first 10 days of simulation. Forecast as well as analysis errors, both for height and for wind components, are reduced by more than a factor of two by application of the bias correction procedure. The climate error in the assimilation is reduced even more. Climate error was

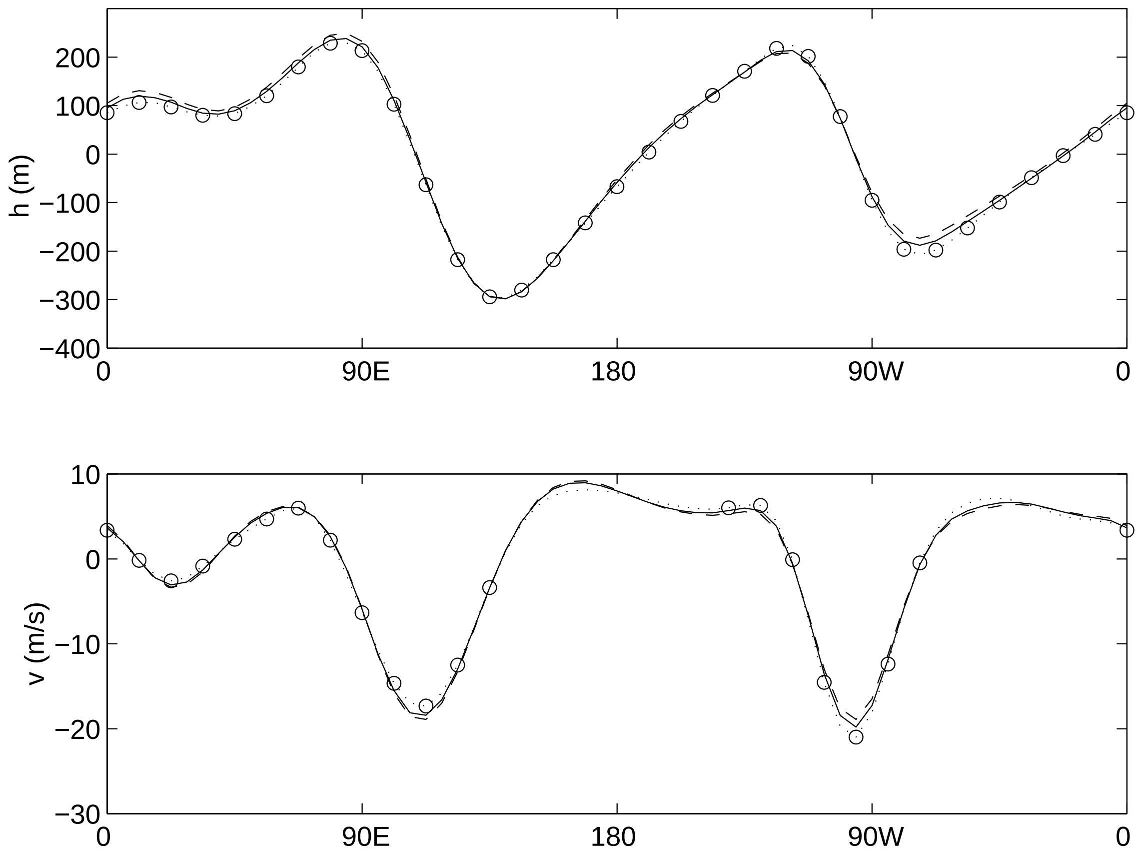

Figure 5: True state (dotted curve), bias-corrected forecast (dashed curve) and bias-corrected analysis (solid curve) after 48 hours of data assimilation. The $h-$component is shown in the top panel; the $v-$component in the bottom panel. Observations are taken every 12 hours as marked.

computed by simulating 40 days of data assimilation, calculating the mean analyzed state over the last 30 days of the assimilation, and comparing the result to the mean true state over the same period. In figure 7 we show the $h-$ and $v-$components of climate error for the assimilation run without bias correction (dashed curve), for the run with bias correction (solid curve), as well as for a climate run of the forecast model in which no data are assimilated (dotted curve). The latter corresponds to the difference between the model and true climates, as can be seen from Figure 1 as well.

The rms climate height error associated with the bias-corrected analyses is $0.85m$, compared to $12.7m$ for the bias-blind analyses and $96.2m$ for the forecast model; the corresponding rms climate wind errors (v-component) are $0.13ms^{-1}, 0.99ms^{-1}$, and $2.9ms^{-1}$, respectively.

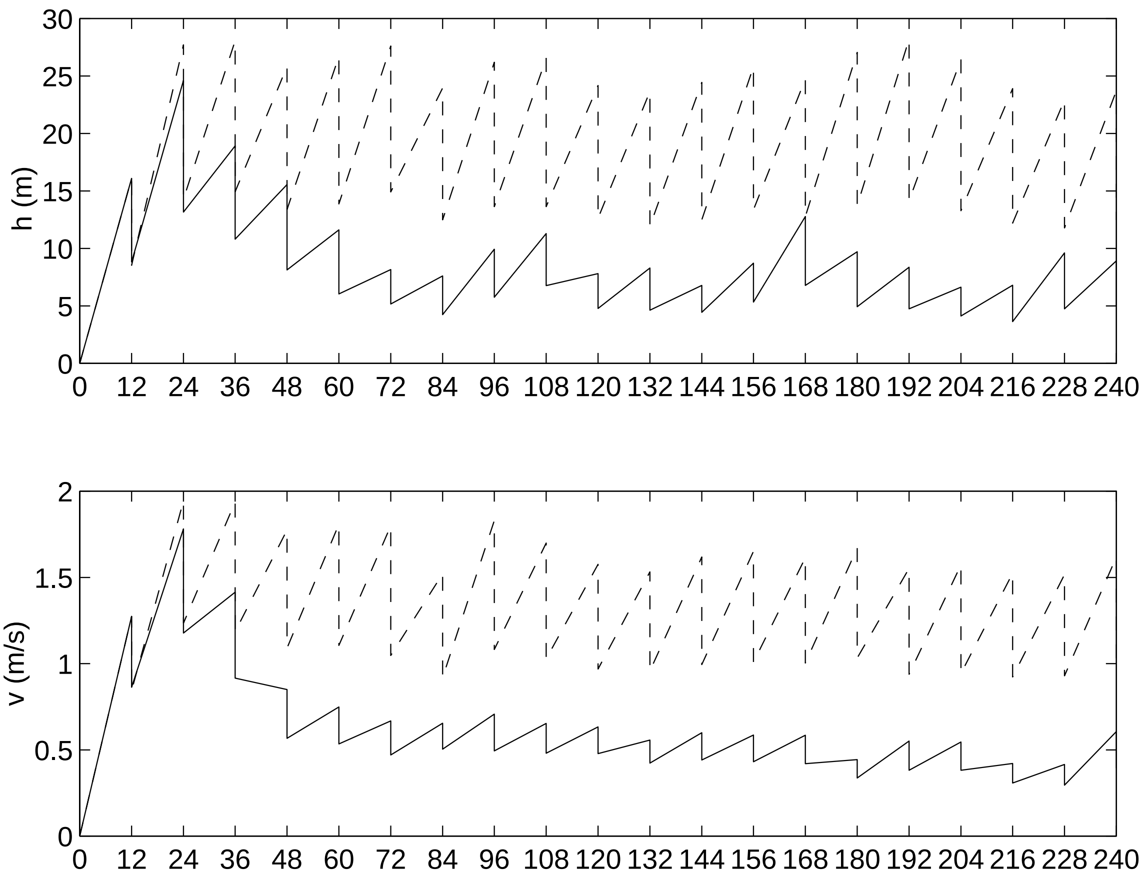

Figure 6: Root-mean square error evolution for the bias-blind estimates (dashed curve) and for the bias-corrected estimates (solid curve). The horizontal axis represents time t_k (in hours).

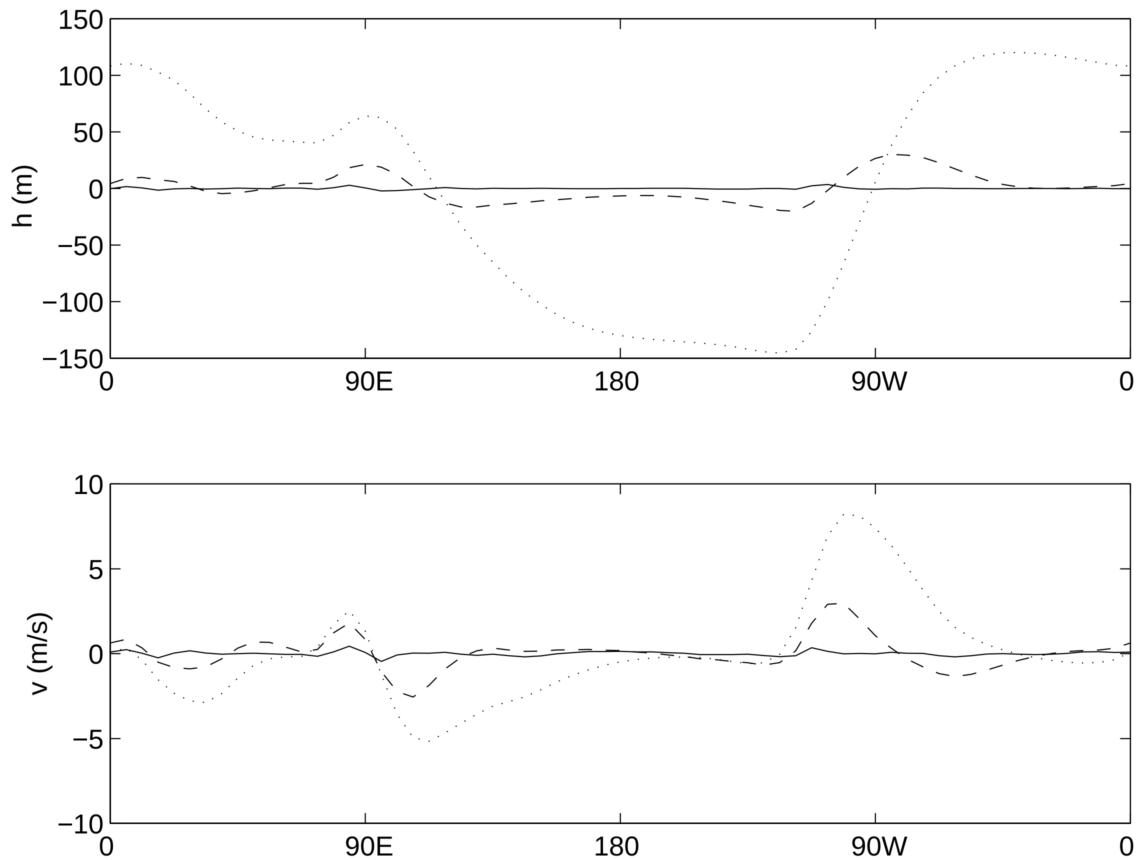

Figure 7: Climate errors computed from 40-day simulations without data assimilation (dotted curve), with data assimilation but without bias correction (dashed curve), and with data assimilation and on-line bias correction (solid curve). The $h-$component is shown in the top panel; the $v-$component in the bottom panel.

6 Conclusion

It is well-recognized in practice that forecast models contain systematic errors which can result in biased forecasts. Forecast bias may have a significant impact on the climatology of assimilated data sets, as well as on the instantaneous accuracy of a particular analysis. The bias problem has been largely ignored, however, in the development of data assimilation methodology. Statistical analysis methods are generally derived under the assumption that forecast errors are strictly random and zero in the mean. In actual implementations of these methods the error statistics are sometimes adjusted in order to reduce analysis bias, but—as we showed in section 2—this is an ad-hoc remedy which is only partially effective, and which causes an increase in the random component of analysis error. The only way to properly account for forecast bias in a statistical analysis scheme is to do so explicitly, by estimating the bias and then correcting the forecast prior to analysis.

We presented a rigorous method for estimating forecast bias in an atmospheric data assimilation system based on an unbiased subset of the available observations. The main components of our sequential bias estimation algorithm are already available in existing statistical analysis systems. The added computational cost of incorporating the algorithm into an operational system is roughly one additional solution of the statistical analysis equation, for a limited number of observations. Off-line forecast bias estimates can be produced as well, using an existing analysis system and stored output from a previous data assimilation run. We plan to perform bias estimation experiments in the immediate future using the Goddard Earth Observing System (Pfaendtner *et al.* 1995).

Forecast bias estimation merely represents the first step in addressing the bias problem as it affects atmospheric data assimilation. Once estimates of forecast bias become available they can be used to remove, or at least reduce, analysis bias. However, the mechanisms responsible for generating forecast bias will cause the model diagnostics to be flawed in many cases. This problem is critical when the output of the system is to be used for climate research, which requires a continuous and unbiased record of prognostic as well as diagnostic atmospheric variables.

Ultimately, routine production and monitoring of forecast bias estimates in an operational data assimilation system should lead to improvements in the formulation of the forecast model itself. In the meantime it will be necessary to consider bias correction methods in which model forcing is continuously being adjusted in order to compensate for the effect of systematic errors in the forecast model. Saha (1992) and Takacs (1996) have developed bias correction schemes for statistical data assimilation, while Derber (1989) and Griffith and Nichols (1996) address the problem in the context of the variational method. In future work we plan to develop bias correction methods which are consistent with the framework provided by the present article.

Acknowledgements. We are grateful to Steve Cohn, Greg Gaspari, Joanna Joiner, Dave Lamich, Richard Ménard, Chris Redder, Leonid Rukhovets, Larry Takacs, Ricardo Todling, and Siegfried Schubert for many interesting and useful discussions about forecast bias. Special thanks to Ricardo Todling for pointing out a number of important references in the literature on separate-bias estimation, and for a careful review of the manuscript.

7 References

Anderson, B. D. O., and J. B. Moore, 1979: *Optimal Filtering.* Prentice-Hall, 357pp.

Baker, N. L., 1991: An adaptive correction procedure for radiosonde geopotential height biases. *Preprints, Ninth Conference on Numerical Weather Prediction,* Denver CO, 192–194.

Bartello, P., and H. L. Mitchell, 1992: A continuous three-dimensional model of short-range forecast error covariances. *Tellus*, **44A**, 217–235.

Cohn, S. E., 1996: Introduction to estimation theory. *J. Met. Soc. Japan*, in press.

Cohn, S. E., and R. Todling, 1996: Approximate data assimilation schemes for stable and unstable dynamics. *J. Met. Soc. Japan*, **74**, 63–75.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457pp.

Dee, D. P., 1991: Simplification of the Kalman filter for meteorological data assimilation. *Q. J. R. Meteorol. Soc.*, **117**, 365–384.

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.*, **123**, 1128–1145.

Derber, J. C., 1989: A variational continuous data assimilation technique. *Mon. Wea. Rev.*, **117**, 2437–2446.

Eyre, J. R., 1992: A bias correction scheme for simulated TOVS brightness temperatures. ECMWF Tech. Memo 176.

Ghil, M., S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981: Applications of estimation theory to numerical weather prediction. Pp. 139–225 in: Bengtsson, L., M. Ghil, and E. Källén, *Dynamic Meteorology: Data Assimilation Methods*, Springer-Verlag, New York, 330pp.

Griffith, A. K., and N. K. Nichols, 1996: Accounting for model error in data assimilation using adjoint methods. Pp. 195–204 in: *Computational Differentiation: Techniques, Applications, and Tools.* Proceedings of the Second International SIAM Workshop on Computational Differentiation, Santa Fe, New Mexico, February 1996. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Friedland, B., 1969: Treatment of bias in recursive filtering. *IEEE Trans. Autom. Contr.*, **AC-14**, 359–367.

Friedland, B., 1978: Notes on separate-bias estimation. *IEEE Trans. Autom. Contr.*, **AC-23**, 735–738.

Hollingsworth, A. and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from rawinsonde data. Part I: The wind field. *Tellus*, **38A**, 111–136.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225–1242.

Ignagni, M. B., 1981: An alternate derivation and extension of Friedland's two-stage Kalman estimator. *IEEE Trans. Autom. Contr.*, **AC-26**, 746–750.

Ignagni, M. B., 1990: Separate-bias Kalman estimator with bias state noise. *IEEE Trans. Autom. Contr.*, **AC-35**, 338–341.

Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory*, Academic Press, New York, 376pp.

Julian, P. R., 1991: RADCOR91 - The new radiosonde radiation error correction procedure. NMC Office Note 374.

Kanamitsu, M., and S. Saha, 1996: Systematic tendency error in budget calculations. *Mon. Wea. Rev.*, **124**, 1145–1160.

Lönnberg, P. and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from rawinsonde data. Part II: The covariance of height and wind errors. *Tellus*, **38A**, 137–161.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.

Mendel, J. M., 1976: Extension of Friedland's bias filtering technique to a class of nonlinear systems. *IEEE Trans. Autom. Contr.*, **AC-21**, 296–298.

Miyakoda, K., and J. Sirutis, 1990: Subgrid scale physics in 1-month forecasts. Part II: Systematic error and blocking forecasts. *Mon. Wea. Rev.*, **118**, 1065–1081.

Pfaendtner, J., S. Bloom, D. Lamich, M. Seablom, M. Sienkiewicz, J. Stobie, and A. da Silva, 1995: Documentation of the Goddard Earth Observing System (GEOS) Data Assimilation System–Version 1. *NASA Tech. Memo. No. 104606*, Vol. **4**, Goddard Space Flight Center, Greenbelt, MD 20771. Available electronically on the World Wide Web as
`ftp://dao.gsfc.nasa.gov/pub/tech_memos/volume_4.ps.Z`

Reynolds, C., R. Gelaro, and T. Murphree, 1996: Observed and simulated Northern

Hemisphere intraseasonal circulation anomalies and the influence of model bias. *Mon. Wea. Rev.*, **124**, 1100–1118.

Richtmyer, R. D., and K. W. Morton, 1967: *Difference Methods for Initial-Value Problems,* 2nd edition. Wiley-Interscience, New York, 405pp.

Rutherford, I., 1972: Data assimilation by statistical interpolation of forecast error fields. *J. Atmos. Sci.*, **29**, 809–815.

Saha, S., 1992: Response of the NMC MRF Model to systematic-error correction within integration. *Mon. Wea. Rev.*, **120**, 345–360.

Schlatter, T. W., 1975: Some experiments with a multivariate statistical objective analysis scheme. *Mon. Wea. Rev.*, **103**, 246–256.

Tacker, E. C., and C. C. Lee, 1972: Linear filtering in the presence of time-varying bias. *IEEE Trans. Autom. Contr.*, **AC-17**, 828–829.

Takacs, L. 1996: A simple bias correction algorithm for use in the GEOS data assimilation system. Submitted to *Mon. Wea. Rev.*

Takacs, L., and M. J. Suarez, 1996: Dynamical aspects of climate simulations using the GEOS General Circulation Model. *NASA Tech. Memo. No. 104606*, Vol. **10**, Goddard Space Flight Center, Greenbelt, MD 20771. Available electronically on the World Wide Web as `ftp://dao.gsfc.nasa.gov/pub/tech_memos/volume_10.ps.Z`

Tenenbaum, J., 1996: Jet stream wind analyses: Comparison of aircraft observations with analyses. *Weather and Forecasting,* **11**, 188–197.

Tibaldi, S., and F. Molteni, 1990: On the operational predictability of blocking. *Tellus*, **42A**, 343–365.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at the National Meteorological Center: The generation of perturbations. *Bull. Amer. Meteorol. Soc.,* **74**, 2317–2330.

Zhou, D. H., Y. X. Sun, Y. G. Xi, and Z. J. Zhang, 1993: Extension of Friedland's separate-bias estimation to randomly time-varying bias of nonlinear systems. *IEEE Trans. Autom. Contr.*, **AC-38**, 1270–1273.